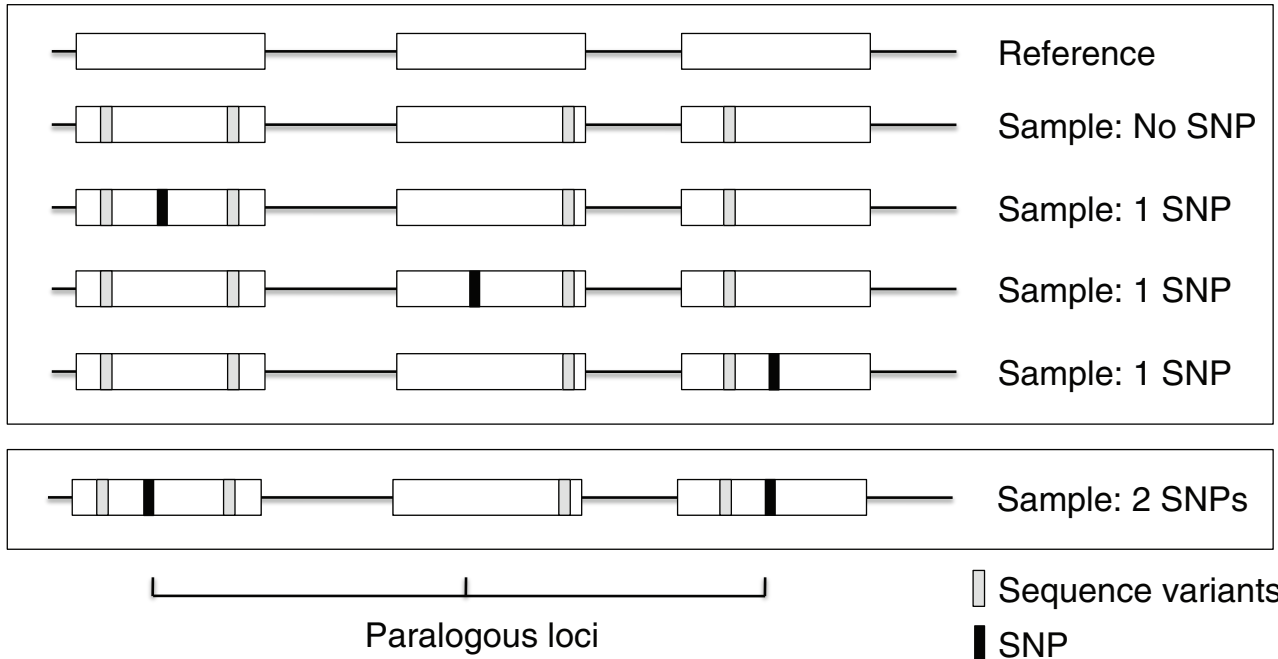


(a) The problem of paralogous SNP identification

Example chromosome



(b) NGS mapping strategies (quality-blind)

For a given read and considering all alignments for that read having $\leq k$ mismatches to a reference template

Unique

- Retain the read if it yields exactly $d=1$ alignment

Best

- Best-guess: Retain the alignment with fewest mismatches (randomly choose among equally best alignments)
- Best-noguess: Retain the alignment with fewest mismatches (discard if multiple equally best alignments exist)

All

- Top-d: Retain up to d alignments
- Max-d: Retain up to d alignments, but discard all alignments if more than d exist